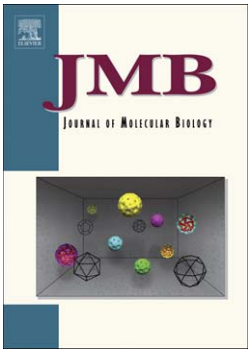


Directed Evolution of *Gloeobacter violaceus* Rhodopsin Spectral Properties

Martin K.M. Engqvist, R. Scott McIsaac, Peter Dollinger, Nicholas C. Flytzanis, Michael Abrams, Stanford Schor, Frances H. Arnold

PII: S0022-2836(14)00308-8
DOI: doi: 10.1016/j.jmb.2014.06.015
Reference: YJMBI 64489

To appear in: *Journal of Molecular Biology*

Received date: 28 April 2014
Revised date: 5 June 2014
Accepted date: 22 June 2014

Please cite this article as: Engqvist, M.K.M., McIsaac, R.S., Dollinger, P., Flytzanis, N.C., Abrams, M., Schor, S. & Arnold, F.H., Directed Evolution of *Gloeobacter violaceus* Rhodopsin Spectral Properties, *Journal of Molecular Biology* (2014), doi: 10.1016/j.jmb.2014.06.015

Directed Evolution of *Gloeobacter violaceus* Rhodopsin Spectral Properties

Martin K. M. Engqvist[1], R. Scott McIsaac[1], Peter Dollinger[1], Nicholas C. Flytzanis[2], Michael Abrams[1], Stanford Schor[1], Frances H. Arnold[1,2*]

1. Division of Chemistry and Chemical Engineering, Mail Code 210-41, California Institute of Technology, Pasadena, California, United States of America
2. Division of Biology and Biological Engineering, Mail Code 156-29, California Institute of Technology, Pasadena, California, United States of America

* To whom correspondence should be addressed. Email: frances@cheme.caltech.edu

Abstract

Proton-pumping rhodopsins (PPRs) are photoactive retinal-binding proteins that transport ions across biological membranes in response to light. These proteins are interesting for light-harvesting applications in bioenergy production, in optogenetics applications in neuroscience, and as fluorescent sensors of membrane potential. Little is known, however, about how the protein sequence determines the considerable variation in spectral properties of PPRs from different biological niches or how to engineer these properties in a given PPR. Here we report a comprehensive study of amino acid substitutions in the retinal binding pocket of *Gloeobacter violacaeus* rhodopsin (GR) that tune its spectral properties. Directed evolution generated 70 GR variants with absorption maxima shifted by up to +/- 80 nm, extending the protein's light absorption significantly beyond the range of known natural PPRs. While proton pumping activity was disrupted in many of the spectrally shifted variants, we identified single tuning mutations that incurrred blue and red shifts of 42 nm and 22 nm, respectively, that did not disrupt proton pumping. Blue-shifting mutations were distributed evenly along the retinal molecule while red-shifting mutations were clustered near the residue K257, which forms a covalent bond with retinal through a Schiff base linkage. Thirty-four of the identified tuning mutations are not found in known microbial rhodopsins. We discovered a subset of red-shifted GRs that exhibit high levels of fluorescence relative to the wild-type protein.

Keywords: optogenetics, retinal, fluorescent proteins, opsins, proton pumps

Abbreviations: bacteriorhodopsin (BR); channelrhodopsin (ChR); *Gloeobacter violaceus* Rhodopsin (GR); *Haloquadratum walsbyi* BR (HwBR); Monterey Bay proteorhodopsin (MBP); *Natronomonas pharaonis* sensory rhodopsin II (NpSRII); proteorhodopsin (PR); proton-pumping rhodopsin (PPR); sensory rhodopsin I (SRI); *Salinibacter ruber* sensory rhodopsin I (SrSRI); sensory rhodopsin II (SRII); quantum mechanics/molecular mechanics (QM/MM)

Introduction

Many proteins, including transcription factors [1, 2], enzymes [3-5], and ion pumps/channels [6-8] have activities that are directly coupled to light. One functionally diverse class of light-activated proteins – opsins – occurs throughout the microbial world, where they facilitate phototactic and photophobic responses in algae [9, 10], enhanced survival in response to nutrient starvation in bacteria [11], and conversion of solar energy into electrochemical energy in diverse microbes [12, 13]. Knowing the molecular determinants for these proteins' light absorption properties will help us to understand their biological functions as well as to engineer new versions for applications in bioenergy [13], biomaterials [14], optogenetics [15], and live-cell imaging [16].

Opsins are integral membrane proteins that bind a retinal chromophore to form a colored holoprotein known as rhodopsin. All microbial rhodopsins have similar overall topologies with seven transmembrane helices that enclose all-*trans* retinal [17, 18], which is bound through a covalent Schiff base linkage to the ε-amino group of a conserved lysine located at helix seven. Despite similar structures, rhodopsins show significant variation in the wavelengths of light they absorb, mediated by interactions between retinal and the opsin apoprotein [19]. Naturally occurring microbial rhodopsins exhibit a range of spectral properties, with peak absorption wavelengths (λ_{max}) between 437 nm and 590 nm [20, 21]. In proteorhodopsins (PR), which have been studied extensively, the distribution of green and blue absorbing pigments is stratified according to ocean depth, with green-absorbing PRs appearing in microbes living near the surface and blue absorbing PRs in deeper water [22]. The visual pigment rhodopsins of cottoid fish endemic to Lake Baikal in eastern Siberia show the same trend, with blue shifts occurring in deeper dwelling species [23, 24]. Because λ_{max} appears to be tuned to the light available in the host's biological niche, nature's many evolution experiments may not have tested the physical-chemical limits of rhodopsin spectral properties. We wish to explore whether rhodopsin light

absorption can be extended beyond the λ_{max} of known homologs and determine the protein sequence changes that mediate such spectral shifts. We also want to know how changes that alter light absorption properties affect other properties such as fluorescence and the protein's ability to pump protons.

Spectral tuning has been reported for proteorhodopsin and bacteriorhodopsin (BR) (both of which are proton pumping rhodopsins (PPRs)), sensory rhodopsins I and II (SRI and SRII), and channelrhodopsins (ChRs), which act as light-activated non-specific cation channels. Random mutagenesis and screening of Monterey Bay proteorhodopsin (MBP) revealed 20 unique color-tuning mutations with 12 red shifts and 8 blue shifts, the majority of which attenuated proton pumping [25]. At pH 4, all red-shifted mutants had λ_{max} within 20 nm of WT MBP, while a single mutation, G138D, incurred a 55 nm blue shift in λ_{max} from 548 nm to 493 nm [25]. Four color-tuning mutations, including G138D, were localized to the putative retinal binding pocket, and roughly half of the color-tuning mutations were found on helices three and four [25]. In *Haloquadratum walsbyi* BR (HwBR), altering the size or polarity of four conserved residues near retinal resulted in small blue shifts (<20 nm) [26]. Accumulation of three of these mutations resulted in a functional HwBR mutant with λ_{max} blue-shifted by ~50 nm [26].

Substituting residues from rhodopsins with different λ_{max} can also result in significant spectral tuning. By replacing single amino acid residues in *Salinibacter ruber* sensory rhodopsin I (SrSRI, λ_{max} = 557 nm) with the corresponding residues from *Natronomonas pharaonis* sensory rhodopsin II (NpSRII, λ_{max} = 498 nm, ~30% amino acid identity to SrSRI), Sudo *et al*. identified three blue tuning mutations, each of which shifted λ_{max} by 14-24 nm [27]. The SrSRI variant with all three mutations had a λ_{max} of 525 nm (32 nm blue shift), accounting for more than half the spectral shift between SrSRI and NpSRII [27]. In ChRs, recombination of variants with different spectral preferences, kinetics, and ion permeability has been used to generate mutants with different combinations of these properties [28], including a highly active red-shifted mutant called ReaChR [29].

Previous work on spectral tuning of rhodopsins has not included a comprehensive investigation of all possible retinal binding pocket mutations. We reasoned that directed evolution could be used to identify mutations in this region, and combinations thereof, that alter λ_{max} and other spectral properties. We chose to focus on *Gloeobacter violacaeus* rhodopsin (GR), a PPR whose strong expression in *Escherichia coli* made it amenable to purification in 96-well plate format for moderate-throughput screening. We targeted 19 residues in the retinal binding pocket for saturation mutagenesis and screened the resulting libraries for shifts in the light absorption spectra. Recombination of the spectral tuning mutations and further directed evolution resulted in GR variants whose λ_{max} shifted by up to +/- 80 nm, the largest spectral shifts of a rhodopsin reported to date. A subset of these mutants exhibits strong fluorescence in the far-red/infrared.

Results

Approach to GR spectral tuning: The retinal binding pocket forms all of the retinal chromophore contacts; we expected that these residues would provide a rich source of spectral tuning mutations. To identify amino acid substitutions in the GR retinal binding that led to shifts in λ_{max}, we developed a moderate-throughput expression and purification assay (Figure 1). Briefly, following GR expression in *E. coli*, protein was extracted from the cell membrane using lysozyme treatment in the presence of the detergent n-dodecyl β-D-maltoside (DDM) and subsequently purified in 96-well filter plates pre-loaded with Ni-NTA resin (Figure 1). The absorption properties of engineered variants were determined in 96-well format using a standard plate reader (Figure 1, see *Materials and Methods* for details). The limits of spectral shifting were further explored by recombination of single mutations identified from saturation mutagenesis libraries and additional mutagenesis and screening of the most red- and blue-shifted recombined variants.

Targeted mutagenesis of GR: To identify putative binding pocket residues in GR, a homology model was generated (see *Materials and Methods*) based on the crystal structure of its XR homolog (44% protein sequence identity). In the GR model, 20 residues (D121, W122, T125, V126, L129, M158, I159, G162, E166, G178, S181, T182, F185, W222, Y225, P226, D253, A256, and K257) are within 5 Å of the retinal chromophore (Figure 2). D121, W222, Y225, P226, D253, and K257 (which forms the Schiff base) are completely conserved between GR and the two best studied PPRs, PR and BR (Table 1, Figure S1). While PPRs show extensive sequence diversity (with overall sequence identities as low as ~20%), the binding pocket residues show a marked degree of conservation (Table 1, Figure S1). Site-saturation mutagenesis was performed at each of the binding pocket positions except K257, which forms the critical retinal linkage. Screening 88 clones for each saturation mutagenesis library (94% library coverage) identified a total of 52 unique blue- and red-shifted variants (Figures 2, 3; Table S1). Mutagenesis at 15 of the 19 binding pocket residues yielded one or more spectral tuning mutations (Figures 2, 3). Thirteen of these 15 positions gave rise to variants with exclusively blue or red shifts, but not both (the exceptions being D121 and A256, which gave both) (Figures 2, 3). 80% of red-shifted variants contained mutations within 4 Å of the Schiff base linkage, whereas blue-shifting mutations were distributed along the retinal chromophore (Figure S2). Blue-shifting mutations far from the Schiff base (>10 Å) tended to increase local hydrophobicity, while blue-shifting mutations closer to the Schiff base tended to decrease local hydrophobicity (Figure 4). The majority of mutations that cause a red shift were identified at D121 and T125, which had 12 and 6 tuning mutations, respectively. Three residues, G162, E166, and A256, each had 4 blue-shifting mutations. Residues G178, W222, and Y225 had a single tuning mutation apiece: G178Q (blue shift), W222M (red shift), and Y225A (red shift). While the majority of binding pocket residues had two or fewer mutations that only modestly affected λ_{max}, a handful of *hot spots* (D121 and G162) had many tuning mutations that led to large shifts in λ_{max}, (Figure 3).

Proton pumping: To determine the ability of the spectrally-shifted GRs to pump protons, we used a modified version of a qualitative, fluorescence-based proton-pumping assay from Kamo *et al.* [30] (detailed in *Materials and Methods*, Figures S3, S4A-C). Briefly, light-driven increases in the proton gradient (via proton pumping by GR) result in efflux of a fluorescent dye from the cell (through multidrug transporters), leading to decreased fluorescence (Figure S4C). The difference in the change of fluorescence between WT GR and each mutant was used to compare their relative pumping capacity. Of the 52 shifted single variants, 32 retained some level of proton-pumping activity (Figure 3), though none appeared to be more active than wild type (data not shown). The most blue-shifted single mutant, G162L (**GRb**lue **1**, GRb1, -42 nm shift), retained proton-pumping capacity (Figures 3, 5). The most red-shifted variant to retain proton-pumping activity was T125N (**GR**red **1**, GRr1, 22 nm shift) (Figures 3, 5). All tuning mutations at D121 and D253 eliminated proton pumping, consistent with studies of BR [31]. Furthermore, all tuning mutations at D121's neighboring residue W122 eliminated proton pumping, while all tuning mutations at L129 and E166 reduced pumping by ~50%. Tuning mutations near the Schiff base were more likely to disrupt proton pumping than tuning mutations far from the Schiff base (Figure 6). While many red- and blue-shifted variants had attenuated pumping capacity, there were exceptions: single mutants G162C, G162V, and G162L are each blue-shifted >30 nm and maintain their ability to pump protons (Figure 7).

Recombination and further mutagenesis: To test whether the tuning mutations could be combined to generate even larger spectral shifts, two recombination libraries were generated, one for blue-shifting mutations (the 'blue' library) and one for red-shifting mutations (the 'red' library). A fractionated plasmid amplification (Figure S5A) approach allowed us to rapidly generate all possible combinations of targeted mutations with high confidence. Mutations that resulted in small spectral shifts and/or eliminated proton pumping were not included in these libraries. Recombining mutations V126A, M158L, G162C, G162L, G162S, G162V, E166P, E166W, S181G, A256C, A256D, A256G and A256S with the wild-type codon for each position

gave a blue library having 600 possible variants (Figure S5B). 1,760 clones were screened (95% coverage). Five mutants were found to be more blue-shifted than GRb1 (G162L). The most blue-shifted variant had three mutations, G162L, E166W, and A256S, and exhibited a further shift of -22 nm with respect to GRb1. This variant is called GRb2 (Figures 3, 5, Table S2).

The red library recombined T125C, T125D, T125G, T125N, T125V, L129K, L129W, W222M, P226I, P226V, and A256M with the wild-type codon for each position (216 possible variants) (Figure S5B). Screening 880 clones (98% coverage) yielded three variants more red-shifted than GRr1 (T125N) (Figure 3, Table S2). The T125C/A256M (GRr2) and T125V/A256M variants were the most red-shifted: each had a ~16 nm additional shift compared to GRr1 (Figures 3, 5, Table S2). None of the eight recombined further blue- or red-shifted variants showed measurable proton-pumping activity (Figure 3).

We wished to explore whether even larger shifts in GR λ_{max} were possible. We therefore performed site-saturation mutagenesis on GRb2 and GRr2 at all sites within the retinal-binding pocket that were not already mutated (excluding K257). For the GRb2 libraries, 1,408 (16 x 88) clones were screened for 94% coverage, and for the GRr2 libraries 1,496 (17 x 88) clones were screened for 94% coverage. Variants further shifted with respect to the GRb2 and GRr2 parents were sequenced. This led to the identification of ten tuning mutations, eight of which had not been identified in the first round of site-saturation mutagenesis (Figure 3, Table S3). The libraries generated with GRr2 yielded two unique mutations that further red-shifted the protein (Figure 3). The most red-shifted variant (GRr3, λ_{max} = 619 nm), with an +81 nm shift relative to WT GR, contained the D121E mutation in addition to the T125C and A225M mutations in GRr2 (Figures 3, 5, Table S3). The spectral shift of GRr3 was greater than predicted by simply adding effects from single tuning mutations (Figure S6). The three most blue-shifted variants, with -79 to -80 nm shifts, had the W122L, W122M or D121E mutations in addition to the parental G162L, E166W, and A256S mutations. The W122L/G162L/E166W/A256S variant is referred to as GRb3 (Figures 3, 5, Table S3). Its λ_{max} is 458 nm. Overall, screening 7,216 clones (1,672 in the first

round of site-saturation mutagenesis, 2,640 for the recombination libraries, and 2,904 in the second round of site-saturation mutagenesis) for shifted absorption maxima resulted in 70 unique shifted variants (Figure 3). GRb3 and GRr3 differ by only seven mutations, yet their λ_{max} are 161 nm apart.

Some GR mutants exhibit increased fluorescence: We tested the fluorescence properties of all 70 spectrally tuned GRs using a 96-well assay developed for this purpose. The full excitation spectrum was computed for each GR variant at multiple emission wavelengths (625 – 775 nm in 25 nm increments). The area under the curve (AUC) of the excitation spectra (at an emission wavelength of 725 nm) for each variant was used to score its overall fluorescence. AUC values were rank-ordered and color-coded according to their absorbance shift, revealing six outlier red-shifted variants (Figure 8A). Measuring the full excitation-emission spectra of these variants confirmed increased fluorescence over the wild type protein, which exhibited no measurable fluorescence under our assay conditions (Figure 8B). The variant displaying the greatest *in vivo* fluorescence was D121E/T125C/A256M (Figure 8B), which also had the highest quantum yield *in vitro* (1.2×10^{-2}) of the six bright mutants (Table 2). Increasing the external pH from 6 to 11 attenuated fluorescence ~2-fold (Figure 8C).

Discussion

We have developed and implemented a moderate-throughput directed evolution strategy to shift GR λ_{max} by +/- 80 nm. We found that the GR retinal-binding pocket contains a rich set of tuning mutations, many of which are not found in known microbial rhodopsins. Collectively, 70 unique mutations shifted λ_{max} by at least 5 nm, and 19 single mutations were identified that shifted λ_{max} by at least 20 nm. While directed evolution was efficient at identifying GR tuning mutations, many of these mutations had deleterious effects on proton pumping. In all tested cases, accumulation of tuning mutations for larger spectral shifts eliminated proton pumping.

Routes to spectral tuning in opsins: The absorption maximum, λ_{max}, is the wavelength

at which the photon energy corresponds to the most probable transition between the resting state (S0) and the excited state (S1) of the retinal chromophore [32, 33]. To achieve blue absorbance shifts (for shorter wavelength, higher energy light), the S0-S1 energy gap must increase, whereas red shifts (for longer wavelength, lower energy light) are achieved by decreasing the S0-S1 energy gap. The stabilization and destabilization of the S0 and S1 states is in turn governed by interactions between retinal and the protein binding pocket. Four factors have been reported to affect the S0-S1 energy gap: (i) conjugation of the retinal chromophore [26, 34], (ii) polarity of the retinal-binding cavity [35, 36], (iii) chromophore positioning effects [37], and (iv) interaction between Schiff base and the Schiff base counterion [38, 39]. In our dataset, blue-shifting mutations are correlated with changes in polarity along the retinal molecule (Figure 4). We speculate that the majority of red-shifting mutations result from disrupting the interaction between the protonated Schiff base and its counterion (thereby also crippling pumping capacity).

GR red shifts: In GR, K257 forms the Schiff base linkage to retinal and D121 serves as counterion to the protonated Schiff base. Based on structural studies from BR, light-induced retinal isomerization reverses the pKa of K257 and D121, resulting in proton transfer from K257 to D121, which is essential for proton pumping [40]. The distribution of red-shifting mutations was strongly skewed, with a total of 18 tuning mutations at D121 and T125 (Figure 4). While the red-shifting mutations T125C, T125D, T125N, T125V, L129K, L129W, W222M, P226I, P226V, and A256M did not disrupt proton pumping, accumulation of these mutations for larger red shifts eliminated proton pumping. Each of these recombined variants had multiple mutations within 5 Å of K257, increasing the likelihood that they would disrupt the Schiff base counterion. Thus, while large red shifts in GR's λ_{max} were achieved, it may not be possible to engineer strongly red-shifted GR variants that also retain proton-pumping activity by mutagenesis at the retinal-binding pocket sites we tested. Tuning mutations outside the binding pocket can be identified by screening libraries made by error-prone PCR mutagenesis [25]; it may be that such mutations

have less deleterious effects on proton pumping. Since natural rhodopsins appear to be tuned to use light in their local environments [22-24], and thermal radiation in deep-sea vents peaks in the infrared [41], we speculate that the microbiota of these deep-sea environments may contain functional PPRs with red-shifted λ_{max}.

Residues D121, W222, Y225, P226, and D253 are completely conserved in the GR, BR and PR retinal binding pockets, but can be mutated to alter spectral properties (Figure 3, Table 1). All but one tuning mutation – D121E – at these residues result in red shifts (Figure 3); however, the absence of these tuning mutations in the natural variants is suggestive of functional constraints at these residues (Table 1). In BR, residues D85 and D212 (corresponding to D121 and D253 in GR) are absolutely essential for pumping [31]. That all tuning mutations at these positions eliminate pumping indicates that these residues have equivalent roles in GR proton pumping. It has been suggested that BR(Y185) and BR(P186) (corresponding to Y225 and P226 in GR) act as a hinge for moving helix 6 away from the main body of the protein during the M → N transition in the BR photocycle [42]. BR(Y185) is also hydrogen-bonded to BR(D212)[43]. We find that the GR(Y225A) tuning mutation eliminates pumping, while tuning mutations GR(P226I) and GR(P226V) do not. Disruption of the hydrogen-bonding network required for proton pumping by GR(Y225A), but not by GR(P226I) and GR(P226V), could explain why Y225A interferes with proton pumping. The final red-shifting mutation in our study that eliminated proton pumping was GR(T125L), which is also conserved in BR (PR shows variation at this position). GR(T125) corresponds to BR(T89). BR(T89) and BR(D85) form hydrogen bonds between their carboxylate groups [44] as part of the BR hydrogen-bonding network. Thus, similar to Y225A, disruption of this hydrogen-bonding network in GR could explain why T125L disrupts proton pumping.

GR blue shifts: Blue-shifting mutations increase the photon energy required for retinal to transition between its resting (S0) and its excited (S1) states [32, 33]. Quantum mechanics/molecular mechanics (QM/MM) models predict that changing the polarity of the

retinal binding pocket can increase the S0-S1 energy gap; more specifically, increasing polarity near the Schiff base or reducing polarity in the region surrounding the β-ionine ring should lead to blue shifts [36]. Our collection of blue tuning mutations provides empirical evidence for this theoretical prediction: blue-shifting mutations that decreased hydrophobicity were clustered near the Schiff base, while those that increased hydrophobicity were near the β-ionine ring (Figure 4).

Unlike red-shifting mutations, mutations that caused blue shifts of λ_{max} did not, in general, disrupt GR proton pumping and were evenly distributed along the retinal molecule (Figures 3, S2). With a λ_{max} at 496 nm, G162L was the most blue-shifted GR variant that retained significant proton-pumping activity (Figure 3). Besides D121E, a mutation of the Schiff base counterion, the only other blue-shifting mutations to eliminate proton pumping were at W122 (located directly next to the counterion). Two of the three tuning mutations at W122, W122A and W122G, resulted from replacing a large amino acid side chain with small ones, indicating that these mutations may incur structural changes that disrupt the ability of K257 to protonate D121, thereby eliminating pumping. In general, changes in residue molecular weight do not correlate with changes in λ_{max} in GR (Figure S7). The third tuning mutation, W122R, is similar in size to the parent residue, but it is much more hydrophilic, and almost always carries a positive charge in proteins at physiological pH. We therefore speculate that by changing the local charge distribution, W122R prevents the proton transfer from the Schiff base to D121, which in turn inhibits proton pumping. In all tested cases, accumulation of blue-shifting mutations for larger blue shifts eliminated proton pumping (Figure 3).

Natural variation in opsins: We investigated whether any of the mutations that shift GR λ_{max} are present in naturally occurring GR homologs. We used GR to search the NCBI nucleotide and environmental sample database for natural variants (as described in *Materials and Methods*, Table S4), GR-like natural variants were defined as proteins with >40% amino acid identity to GR. Of the 52 unique spectral tuning mutations uncovered after the first round of site-saturation mutagenesis of GR, only one, S181A (a blue-shifting mutation), is found in a GR-

like natural variant (Table 1). That tuning mutations are not found in GR-like natural variants suggests that either (1) GRs have not evolved for light absorption at different wavelengths or (2) GRs may have evolved alternative ways to harvest different colors of light (discussed below).

We also used BR (21% amino acid identity to GR) and PR (22% amino acid identity to GR) to search for natural variants that are more distant relatives of GR. Collectively, 31 BR-like and 35 PR-like natural variant sequences were found (i.e., sequences with >40% amino acid identity to PR or BR). This set of 66 BR/PR-like natural variants contained five of the GR spectral tuning mutations: V126T, M158A, M158L, G162V and A256S, all five of which caused blue shifts in GR. GR variants with these tuning mutations retained proton-pumping capacity (Figure 3). Thus, tuning mutations that do not disrupt proton-pumping capacity in a PPR may also not disrupt proton pumping in distantly related PPRs due to the marked degree of conservation of the retinal binding pocket, which is supported by results from several rhodopsin engineering studies [27-29].

While known rhodopsins have a wide range of λ_{max} from 437 nm – 590 nm, the variants at both extremes are non-PPRs (Table S5): the range of λ_{max} of reported PPRs is 490 nm – 568 nm (Table S5). The most blue-shifted functional GR variant (G162L) was within this range with λ_{max} = 496 nm, making it more blue-shifted than all reported PPRs except blue proteorhodopsin (BPR), originally isolated from the deep ocean [22]. The most red-shifted GR variant (T125N), with λ_{max} = 560 nm, is well within the range of known PPRs. Collectively, we identified twelve blue-shifted GR variants and four red-shifted with λ_{max} beyond the range of known PPRs, none of which retained proton capacity (Figure 3). The most blue- and red-shifted GR mutants had λ_{max} of 458 nm and 619 nm, respectively. Since extending GR λ_{max} beyond the known natural range of functional PPRs eliminates its proton pumping ability, it is possible that there is a physical limit to the light a rhodopsin can harness through retinal alone and still retain proton pumping, the mechanistic basis of which is currently unclear. Of course, our experiments have only investigated the effects of mutations in the retinal binding pocket and have not explored the

possibly compensatory effects on proton pumping of mutations outside this region.

Salinixanthin, a carotenoid antenna: It has been reported that GR and its close homolog XR bind both retinal and a carotenoid antenna called salinixanthan, which enables the absorption and conversion of blue light into proton pumping activity [45, 46]. The XR crystal structure implicated a particular glycine residue (G162 in XR) as important for salinixanthan binding. Since tryptophan at the homologous position in BR would physically block salinixanthan binding, it was suggested that G162 may be useful as a diagnostic for salinixanthan binding in PPRs [47]. By introducing tryptophan into the homologous GR position (G178), salinixanthan binding was indeed eliminated [45].

Thirty of the 32 GR-like homologs (i.e., sequences with >40% amino acid identity to GR) contained G178, indicating they could be salinixanthan binders. Two of the 32 natural GR variants have large hydrophobic residues at G178 (G178F in *Isosphaera pallida* [gi: 319748837] and G178W from a metagenome library [gi: 129721342]), suggesting the existence of natural GR-like variants unable to bind salinixanthan.

We found only one of the GR spectral tuning mutations (S181A) in natural GR-like variants. This may partially be due to GR's ability to harvest blue light through salinixanthin, which would have removed selection pressure for finding blue-shifting mutations. In contrast, five GR spectral tuning mutations (all blue-shifts) were found in BR- and/or PR-like natural variants (Table 1). The enrichment for blue-shifting mutations in BR- and PR-like natural variants may be due to their use of tuning mutations to absorb blue-shifted light through the retinal chromophore without an additional carotenoid antenna.

Comparison to previously engineered opsins: To investigate whether the spectral-shifting mutations reported here had been described previously, we surveyed the literature on rhodopsin spectral shifts resulting from mutations introduced into the retinal binding pockets of the GR homologs BR and PR (Table 1). Sixteen of the 52 GR mutations reported here have been described in engineered PR or BR variants (Table 1, underlined). Twelve of these gave the

same type of spectral shift in GR as in PR or BR, whereas four generated the opposite shifts, indicating that interactions between retinal and the opsin apoprotein that determine λ_{max} will not always be conserved due to structural differences between these proteins. A whole-protein random mutagenesis screen for changes in PR spectral properties by Kim *et al.* isolated two blue-tuning mutations, PR(W98G) and PR(V102A), that we also identified as blue-tuning mutations in GR (Table 1) [25]. Overall, our GR dataset contains 34 tuning mutations that are completely new to proton-pumping rhodopsins.

Fluorescent GRs: Archaerhopsin-3 (Arch), a BR-like homolog from *Halorubrum sodomense* (Figure S1), and PR have been reported to exhibit dim fluorescence that responds to changes in membrane voltage when expressed heterologously in mammalian neurons (Arch) and *E. coli* (PR) [16, 48]. The absence of proton pumping is desirable for rhodopsin-based biological sensors that monitor membrane potential without perturbing it in response to light [16, 48]. We thus tested the ability of all 70 engineered GRs to fluoresce across a range of excitations. Collectively, we identified six GR variants that fluoresced more strongly than wild type and the other variants (Figures 8A/8B). None of these fluorescent membrane proteins had proton-pumping activity. All of the strongly fluorescent GR proteins were mutated around the retinal Schiff-base and had a red-shifted λ_{max} compared to WT GR (Figure 3, Figure 8A). The most brightly fluorescing mutant was D121E/T125C/A256M (8B, Table 2), which was also the most red-shifted GR protein engineered in this study (Figure 3, λ_{max} = 619). However, there is not a one-to-one correspondence between the magnitude of red spectral shifts and overall levels of fluorescence.

Studies in bovine rhodopsin have shown that the retinal excited state, which is formed by the absorption of a photon by the retinal ground state, efficiently isomerizes if there is no energy barrier on the path between these two states [49, 50]. Artificially imposing such an energy barrier by restraining the photo-isomerization of retinal extends the excited state lifetime [51]. We speculate that the mutations in the six fluorescent GR proteins create such an energy

barrier on the retinal excited state, thus extending its lifetime. This, in turn, decreases the probability of entering the photocycle (thus inhibiting pumping) while the absorbed energy is emitted as fluorescence. The identification of mutations that significantly improve GR fluorescence provides a first step towards the development of brighter opsin-based biological sensors.

Materials and Methods

Plasmids and bacterial strains

The gene coding for GR was obtained on the GR_pkj900 plasmid from Janos Lanayi. GR was amplified by PCR using primers that excluded the stop codon (Table S6, primers 1-2). The PCR product was cloned by one-step isothermal assembly [73] into an *Nde*I and *Not*I digested pET21a expression plasmid (EMD Millipore, Darmstadt, Germany), adding a flexible five amino-acid-long C-terminal linker and a His_6-tag. This generated the pETME10 (GR) expression plasmid. To construct a translational fusion of GR with a C-terminus CFP the pETME10 plasmid was linearized with XhoI and subsequently purified using agarose gel. CFP was amplified by PCR using primers with an overhang to the linearized plasmid (Table S6, primers 75-76). DNA fragments were then assembled using one-step isothermal DNA-assembly. All constructs were verified by sequencing (Laragen, Culver City, CA, USA) using T7-specific primers (Table S6, primers 77-78). Sequences for the finished His-tagged GR protein and GR-CFP fusion construct can be found in the supplement. All cloning was done in the *Escherichia coli* strain DH5α. The *E. coli* strain BL-21 (DE 3) was used for expression in all experiments.

Homology models of GR

A GR homology model was produced based on the xanthorhodopsin (XR) crystal structure (PDB ID: #3DDL, 44% protein sequence identity) obtained from The Protein Data Bank (PDB, http://www.rcsb.org/pdb/). This model was used to identify amino acids lining the GR retinal binding pocket (those *within* 5 Å of bound retinal). These comprise D121, W122, T125, V126, L129, M158, I159, G162, E166, G178, S181, T182, F185, W222, Y225, P226, D253, A256 and K257. The homology model was generated using the SWISS-MODEL workspace (http://swissmodel.expasy.org/workspace/) [74, 75].

Generation and screening of site-saturation libraries

Site-saturation mutagenesis libraries were constructed by amplification of pETME10 (and subsequent derivatives) using mutagenic primers in a modified QuickChange protocol (Stratagene, La Jolla, CA). Each primer pair had a 20 base 5'-end complementary sequence – to enable recirculation of the amplified, linear DNA – and a 3' non-overlapping sequence designed to have a melting temperature (T_m) of approximately 55°C. This 3' sequence allows for primer binding to the PCR product and allows for exponential amplification of mutated DNA, in contrast with the classical QuickChange protocol. For GR site-saturation libraries, the forward primer additionally carried the degenerate base triplet NNK in between the 5' and 3' sequences (Table S6, primers 3-40). For the second round of site-saturation mutagenesis on GR the same primers were used except in a few cases where the primers overlapped with mutated sites and special primer pairs had to be made (Table S6, primers 65-72 for red-shifted variant and primers 73-74 for the blue-shifted variant). The mutagenic plasmid amplifications were performed using the Phusion High Fidelity PCR Kit (cat. no. E0553S, New England Biolabs, Ipswich, MA) according to the manufacturer's instructions, but with reduced annealing temperature (53°C-54°C). Following PCR, the reaction mixtures were incubated with 0.5-1 µL *Dpn*I (cat. no. R0176S, New England Biolabs, Ipswich, MA) for 1h at 37°C to degrade template DNA. PCR products were then purified over agarose gel electrophoresis and gel extracted by using the QIAquick® gel extraction kit (cat. no. 28704, QIAGEN, Venlo, Netherlands). 5 µL of purified PCR products were used for the one-step isothermal DNA-assembly reaction developed by Gibson *et al.* [73] in order to recircularize the amplified vector. ~1 µL of the reaction mixtures were used to transform electrocompetent BL21 (DE3) by electroporation, which were subsequently plated on selective media. The individual libraries were cloned into *E. coli* and screened to 94% coverage (88 clones at each site) using a high-throughput assay designed to determine the λ_{max} of each rhodopsin variant (see below). Percent coverage was estimated according to $P_i = 100(1-\exp(-T \times F_i))$, where P_i is the probability that a particular sequence i is found in the library, T is the number of clones to screen, and F_i is the frequency with which sequence i occurs in the library [76]. In the case of an NNK library, F_i = ½ x ½ x ¼ = 1/32, which requires $T = 88$ clones to screen to P_i = 94% coverage. For each site-saturation library, ~10 clones were sequenced to confirm the library quality. For example, in the D121 NNK library, we identified three of G, two instances of S, and one instance each of E, L, P, T and V at D121. The annealing temperature of PCR reactions was reduced at least 5 degrees below the wild type primer Tm – calculated while excluding the three bases targeted by NNK – in part to account for differences in annealing temperatures for different NNK sequences.

Generation of recombination libraries

The red recombination library was designed by combining all red shifting mutations that did not destroy the proton pumping ability. This yielded a library of 216 variants (obtained by multiplying the possible codons at each mutagenized site, Figure S5B). If the same criteria were used to make the blue shifted recombination library it would result in a library of over 9000 variants. This number was too large to screen, and the blue shifted recombination library was therefore reduced by setting a threshold for clones with an absorption maximum below 528 nm for sites with more than one mutation. This reduced the library size from 9000 to 600 variants (Figure S5B). Recombination libraries were constructed by amplifying the gene in fragments such that all mutation sites were either covered by mutagenic primers or included by mixing mutated backbones as templates in PCR. The relative binding site of primers used to generate each fragment are shown in Figure S5A). Primers are listed in Table S6. For each site where multiple primers are listed, these were mixed at ratios resulting in even generation of all desired codons. For the red library, primers 43-48 and 49-52 were used to mutate T125, L129, W222 and P226 in one PCR product. Primers 53 and 54 were used to amplify the plasmid backbone from a mix of plasmids containing the wild-type codon at A256 and the codon for A256M. For the blue library, primers 62 and 63-64 were used to generate a PCR product where sites S181 and A256 were mutated. Primers 55 and 56-61 were used to amplify the plasmid backbone while at the same time mutating sites M158, G162 and E166 from a mix of plasmid containing the wild-type codon at V126 and the codon for V126A. The PCR fragments were ligated using one-step isothermal assembly and transformed into *E. coli*.

Expression of libraries

Single *E. coli* BL21 (DE3) colonies transformed with library plasmids or control plasmid (parent) were picked to inoculate 200 µL LB media containing 100 µg/mL of ampicillin (Amp) in 96-deep-well plates with 2 mL well volume (Greiner Bio-One, Kremsmünster, Austria). The plates were covered with a microporous membrane (Easy App™, cat. no. 2978-5827, USA Scientific, Ocala, FL) and incubated in a rotary shaker at 225 rpm, 37°C and 80% humidity. The overnight cultures were diluted 1:20 in LB-Amp media to a final volume of 1 mL in 96-deep-well plates and further incubated 225 rpm, 30°C and 80% humidity for 2 hours. Expression was induced through the addition of 5 µL of 100 mM IPTG and 2 mM retinal in 95% EtOH/5% H_2O to a final concentration of 0.5 mM IPTG and 10 µM retinal. Protein expression proceeded for 4 h at 225 rpm, 30°C and 80% humidity in the dark. The *E. coli* cells were then collected by centrifugation at 4500 RCF for 10 minutes in a swing-out centrifuge (Allegra 25R, Beckman Coulter, Brea, CA).

Method for screening GR spectral tuning libraries

To measure the absorption spectrum the GR protein first has to be extracted from the *E. coli* membrane. Following library expression (see above), cell pellets were resuspended in 100 µL extraction buffer (20 mM Tris-HCl, 200 mM NaCl, 0.15% (w/v) *n*-dodecyl β-D-maltoside, pH 7.5) and frozen at -20°C overnight in order to improve cell lysis. The pellets were thawed at room temperature. 100 µL extraction buffer containing 1.4 mg/mL lysozyme and a small amount of DNaseI was added to the wells. Extraction was allowed to proceed at room temperature for 30 min in the dark. Cell debris was subsequently pelleted by centrifugation at 4500 RCF for 10 minutes. Rhodopsins were purified from the supernatant using Ni-NTA affinity chromatography in a 96-well plate format. For this purpose, each well of a 96-well filter plate (Acroprep™, 1 mL, 1.0 µm glass fiber, Pall Corp., Port Washington, NY) was filled with 125 µL Ni-NTA agarose beads (cat. no. 30230, QIAGEN). The rhodopsin-containing supernatant was transferred to the wells and incubated for 5 min at room temperature followed by a single washing step with 400 µL wash buffer (20 mM Tris-HCl, 200 mM NaCl, 0.15% (w/v) *n*-dodecyl β-D-maltoside, 80 mM imidazole, pH 7.5, incubation 1 min) and one elution step (200 µL elution buffer (20 mM Tris-HCl, 200 mM NaCl, 0.15% (w/v) *n*-dodecyl β-D-maltoside, 400 mM imidazole, pH 7.5, incubation time 5 min). After each incubation step, liquids were forced through the filters by centrifugation (200 RCF, 2 min). Absorption spectra of the purified rhodopsins were recorded in a 96-well microtiter plate from 420 nm to 680 nm in 3 nm steps with 5 flashes per measurement (infinite M200, Tecan).

Analysis of spectral tuning data

Data from the spectral tuning libraries were analyzed and processed using Excel (Microsoft) as well as custom software in python (www.python.org). A nine data-point smoothing function was applied to all measured spectra to decrease the effect of noise in the measurements. The λ_{max} of each clone was then computed and those clones with a λ_{max} that was more than 5 nm shifted away from the parent protein and had an absorption level that was significantly above background were re-screened in quadruplicate and subsequently sequenced.

Fluorescence screening

Following library expression (see above), cell pellets were re-suspended in 700 µL of 200 mM NaCl. Transparent 96-well, flat-bottom plates were pre-aliquoted with 20 µL of 500 mM potassium phosphate buffer at pH 7. A multi-channel pipette was used to transfer 180 µL of cells

to the buffer-containing wells, for a final concentration of 180 mM NaCl and 50 mM potassium phosphate buffer. Flu orescence measurements were performed using a Tecan plate reader, with the instrument mode was set to 'Fluorescence Top Reading' with multiple reads per well (2x2 square). Other instrument parameters included manually setting the gain (100), number of flashes (25), and integration time (20 µs). In Figure 8C, GR(D121E/T125C/A256M) is fused to CFP and the *Normalized Flourescence* is defined as A*(Opsin Fluorescence/CFP Fluorescence) where A is an arbitrary scalar.

Quantum yield determination

GR variants were expressed in 500 mL *E. coli* cultures. Detergent solubilized lysates were loaded onto 1 mL His-tag purification columns and purified using an ÄKTAxpress (GE Healthcare, Little Chalfont, UK) chromatography system into buffer containing 20 mM Tris, 200 mM NaCl, and 0.15% DDM at pH 7.5. Purified proteins were buffer exchanged into 10 mM Tris, 200 mM NaCl, and 0.15% DDM at pH 6.5. ~0.5 mg of Alexa Fluor 680 carboxylic acid, succinimidyl ester (part # A-20008, Life Technologies, Carlsbad, CA) was diluted into phosphate buffered saline (PBS, pH 7.4). Protein and dye samples were diluted to an A_{600} of 0.04. Emission spectra were measured between wavelengths of 615 nm and 850 nm using a Tecan plate reader following excitation at a wavelength of 600 nm. The quantum yields of GR variants (QY_{GR}) were calculated according to the formula $QY_{GR} = QY_{Alexa}(n_{GR}/n_{Alexa})^2(A_{Alexa}/A_{GR})(I_{GR}/I_{Alexa})$. The quantum yield of Alexa Fluor 680 (QY_{Alexa}) is 0.36. Since PBS and protein buffer both have refractive indices of ~1.33, $(n_{GR}/n_{Alexa})^2 \sim 1$. Since the absorbance of the protein and dye samples at a wavelength of 600 nm was 0.04, $A_{Alexa}/A_{GR} = 1$. Therefore, $QY_{GR} = 0.36(I_{GR}/I_{Alexa})$, where I_{GR} and I_{Alexa} are the integrate emission spectra of the protein and dye, respectively. Emission spectra were fit separately to 3rd, 4th, and 5th order Gaussian curves in MATLAB (The Mathworks, Inc., Natick, MA), and then integrated and averaged to estimate I_{GR} and I_{Alexa}. For each variant, we measured the emission spectrum independently three times.

Proton pumping assay

Proton-pumping ability of engineered GRs was measured in a protocol adapted and modified from Kamo *et al.* [30]. The assay relies on the DNA-intercalating dye 2′-(4-ethoxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5′-bi-1H-benzimidazole (H33342) as a fluorescent reporter. The dye fluoresces strongly when inside the cell and less strongly in the extracellular medium. Furthermore, *E. coli* has multi-drug efflux proteins that use the energy

stored in the cellular proton gradient to export H33342 from the cell and thereby cause a decrease in fluorescence (Figure S3). Conceptually, the assay involves starving rhodopsin-containing *E. coli* cells to remove the intrinsic proton gradient and then re-establishing it through light-activated proton-pumping by rhodopsins. To perform the proton-pumping assay, rhodopsin variants were expressed as described above. After 2.5 h expression, 200 µL of each culture were transferred to a clean 96 deep-well plate. Cells were washed with 300 µL phosphate buffer (1 mM KH_2PO_4, 50 mM KCl, pH 7.5) and centrifuged at 5000 rcf for 5 min. Following this wash step, cells were re-suspended in 300 µL phosphate buffer and 1 mM H33342. Plates were incubated with shaking in the dark at 30°C for at least 2.5 h to starve the cells and remove the intrinsic proton-pumping gradient. 200 µL of cells were transferred to a microtiter plate. Fluorescence was measured from the top of the wells with 5 flashes per measurement and 20 µs integration time in a Tecan plate reader. The plate reader temperature was fixed at 30ºC. Emission was measured at 460 nm with excitation at 346 nm until fluorescence reached a steady state, at which point the plate was removed from the Tecan. Light was then applied to the samples for 25 min from a 250 W halogen bulb at a distance of 20 cm with an IR/UV-cut-off filter (Hot-Mirror HMC-1033, UQG Optics) and a 1.5 cm water layer in between to remove UV and IR light (Figures S4A, S4B). The light intensity reaching the microtiter plate was about 3400 µmol $m^{-2}s^{-1}$. Decreased fluorescence, which is a measure of active proton pumping, was then measured in the Tecan plate reader. For positive controls, 4 µL of 2 M glucose was used to cause light-independent export of H33342. For negative controls the non-pumping GR(D121N) mutant was used (Figure S4C). The fluorescence values for each well were normalized to the steady-state fluorescence reached prior to illumination, and the ability of mutants to pump protons was estimated as the difference in fluorescence intensity before and after illumination. The measured decrease in fluorescence for wild type GR was 0.27 ± 0.05 while the average decrease for the negative control (D121N) was 0.05±0.02. The assay positive controls (the addition 2 M glucose in the absence of light) had average decreases of 0.37±0.02. Protein expression was confirmed by the formation of pigment in each well, but not explicitly quantified. The assay was performed independently three times for each GR variant.

Analysis of natural sequence diversity in retinal binding pocket

We investigated whether any of the mutations that shift λ_{max} of GR are present in other known naturally occurring proton-pumping opsins. Bacteriorhodopsins and proteorhodopsins have very low sequence identity to GR (BR, 21% identity and PR, 22% identity) and we therefore first had

to find which amino acid positions in BR and PR correspond to the 20 amino acids in the retinal binding pocket of GR. A structure-guided alignment of the GR, BR and PR retinal binding pockets was made by aligning the GR homology model with a BR crystal structure (1C3W) and PR NMR structure (2L6X) in PyMOL and manually identifying the amino acids corresponding to GR positions D121, W122, T125, V126, L129, M158, I159, G162, E166, G178, S181, T182, F185, W222, Y225, P226, D253, A256, and K257 (Figure S1). We then used the GR (UniProt ID: Q7NP59), BR (UniProt ID: P02945) and PR (UniProt ID: Q9F7P4) protein sequences to search for similar sequences in the NCBI nucleotide and environmental sample databases using the basic local alignment search tool (BLAST). This resulted in a set of 32 sequences for GR, 31 sequences for BR and 35 sequences for PR. From this set of sequences the natural variation at each site of interest was determined using custom software in Python.

Acknowledgments

We thank Janos Lanyi for kindly providing us with the GR gene used in this project, Naoki Kamo for helpful comments on the proton-pumping assay, and Timothy Wannier for critically reviewing our manuscript and assisting with quantum yield measurements. MKM. E. received funding from the German Research Foundation (DFG) under program EN 957/1-1. Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health (R21MH103824) and by the Institute of Collaborative Biotechnologies (contract W911NF-09-D-0001) through the U.S. Army Research Office. The content is solely the responsibility of the authors and does not represent the official views of any of the funding agencies.

References

[1] Masuda S, Bauer CE. AppA is a blue light photoreceptor that antirepresses photosynthesis

gene expression in Rhodobacter sphaeroides. Cell. 2002;110:613-23.

[2] Liu H, Yu X, Li K, Klejnot J, Yang H, Lisiero D, et al. Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. Science. 2008;322:1535-9.

[3] Stierl M, Stumpf P, Udwari D, Gueta R, Hagedorn R, Losi A, et al. Light modulation of cellular cAMP by a small bacterial photoactivated adenylyl cyclase, bPAC, of the soil bacterium Beggiatoa. The Journal of Biological Chemistry. 2011;286:1181-8.

[4] Iseki M, Matsunaga S, Murakami A, Ohno K, Shiga K, Yoshida K, et al. A blue-light-activated adenylyl cyclase mediates photoavoidance in Euglena gracilis. Nature. 2002;415:1047-51.

[5] Ryu MH, Moskvin OV, Siltberg-Liberles J, Gomelsky M. Natural and engineered photoactivated nucleotidyl cyclases for optogenetic applications. The Journal of Biological Chemistry. 2010;285:41501-8.

[6] Lozier RH, Bogomolni RA, Stoeckenius W. Bacteriorhodopsin: a light-driven proton pump in Halobacterium Halobium. Biophysical Journal. 1975;15:955-62.

[7] Nagel G, Ollig D, Fuhrmann M, Kateriya S, Musti AM, Bamberg E, et al. Channelrhodopsin-1: a light-gated proton channel in green algae. Science. 2002;296:2395-8.

[8] Nagel G, Szellas T, Huhn W, Kateriya S, Adeishvili N, Berthold P, et al. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. Proceedings of the National Academy of Sciences of the United States of America. 2003;100:13940-5.

[9] Berthold P, Tsunoda SP, Ernst OP, Mages W, Gradmann D, Hegemann P. Channelrhodopsin-1 initiates phototaxis and photophobic responses in chlamydomonas by immediate light-induced depolarization. The Plant Cell. 2008;20:1665-77.

[10] Sineshchekov OA, Jung KH, Spudich JL. Two rhodopsins mediate phototaxis to low- and high-intensity light in Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences of the United States of America. 2002;99:8689-94.

[11] Gomez-Consarnau L, Akram N, Lindell K, Pedersen A, Neutze R, Milton DL, et al. Proteorhodopsin phototrophy promotes survival of marine bacteria during starvation. PLoS Biology. 2010;8:e1000358.

[12] Claassens NJ, Volpers M, dos Santos VA, van der Oost J, de Vos WM. Potential of proton-pumping rhodopsins: engineering photosystems into microorganisms. Trends in Biotechnology. 2013;31:633-42.

[13] Walter JM, Greenfield D, Liphardt J. Potential of light-harvesting proton pumps for bioenergy applications. Current Opinion in Biotechnology. 2010;21:265-70.

[14] Wagner NL, Greco JA, Ranaghan MJ, Birge RR. Directed evolution of bacteriorhodopsin for applications in bioelectronics. Journal of the Royal Society, Interface / the Royal Society. 2013;10:20130197.

[15] Toettcher JE, Voigt CA, Weiner OD, Lim WA. The promise of optogenetics in cell biology: interrogating molecular circuits in space and time. Nature Methods. 2011;8:35-8.

[16] Kralj JM, Hochbaum DR, Douglass AD, Cohen AE. Electrical spiking in Escherichia coli probed with a fluorescent voltage-indicating protein. Science. 2011;333:345-8.

[17] Spudich JL, Yang CS, Jung KH, Spudich EN. Retinylidene proteins: structures and functions from archaea to humans. Annual Review of Cell and Developmental Biology. 2000;16:365-92.

[18] Zhang F, Vierock J, Yizhar O, Fenno LE, Tsunoda S, Kianianmomeni A, et al. The microbial opsin family of optogenetic tools. Cell. 2011;147:1446-57.

[19] Makino CL, Kraft TW, Mathies RA, Lugtenburg J, Miley ME, van der Steen R, et al. Effects of modified chromophores on the spectral sensitivity of salamander, squirrel and macaque cones. The Journal of Physiology. 1990;424:545-60.

[20] Klapoetke NC, Murata Y, Kim SS, Pulver SR, Birdsey-Benson A, Cho YK, et al. Independent optical excitation of distinct neural populations. Nature Methods. 2014;11:338-46.

[21] Govorunova EG, Sineshchekov OA, Li H, Janz R, Spudich JL. Characterization of a highly efficient blue-shifted channelrhodopsin from the marine alga Platymonas subcordiformis. The Journal of Biological Chemistry. 2013;288:29911-22.

[22] Beja O, Spudich EN, Spudich JL, Leclerc M, DeLong EF. Proteorhodopsin phototrophy in the ocean. Nature. 2001;411:786-9.

[23] Bowmaker JK, Govardovskii VI, Shukolyukov SA, Zueva LV, Hunt DM, Sideleva VG, et al. Visual pigments and the photic environment: the cottoid fish of Lake Baikal. Vision Research. 1994;34:591-605.

[24] Hunt DM, Fitzgibbon J, Slobodyanyuk SJ, Bowmaker JK. Spectral tuning and molecular evolution of rod visual pigments in the species flock of cottoid fish in Lake Baikal. Vision Research. 1996;36:1217-24.

[25] Kim SY, Waschuk SA, Brown LS, Jung KH. Screening and characterization of proteorhodopsin color-tuning mutations in Escherichia coli with endogenous retinal synthesis. Biochimica et Biophysica Acta. 2008;1777:504-13.

[26] Sudo Y, Okazaki A, Ono H, Yagasaki J, Sugo S, Kamiya M, et al. A blue-shifted light-driven proton pump for neural silencing. The Journal of Biological Chemistry. 2013;288:20624-32.

[27] Sudo Y, Yuasa Y, Shibata J, Suzuki D, Homma M. Spectral tuning in sensory rhodopsin I from Salinibacter ruber. The Journal of Biological Chemistry. 2011;286:11328-36.

[28] Wang H, Sugiyama Y, Hikima T, Sugano E, Tomita H, Takahashi T, et al. Molecular determinants differentiating photocurrent properties of two channelrhodopsins from chlamydomonas. The Journal of Biological Chemistry. 2009;284:5685-96.

[29] Lin JY, Knutsen PM, Muller A, Kleinfeld D, Tsien RY. ReaChR: a red-shifted variant of channelrhodopsin enables deep transcranial optogenetic excitation. Nature Neuroscience. 2013;16:1499-508.

[30] Kamo N, Hashiba T, Kikukawa T, Araiso T, Ihara K, Nara T. A light-driven proton pump from

Haloterrigena turkmenica: functional expression in Escherichia coli membrane and coupling with a H+ co-transporter. Biochemical and Biophysical Research Communications. 2006;341:285-90.

[31] Brown LS, Jung KH. Bacteriorhodopsin-like proteins of eubacteria and fungi: the extent of conservation of the haloarchaeal proton-pumping mechanism. Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology. 2006;5:538-46.

[32] Bravaya K, Bochenkova A, Granovsky A, Nemukhin A. An opsin shift in rhodopsin: retinal S0-S1 excitation in protein, in solution, and in the gas phase. Journal of the American Chemical Society. 2007;129:13035-42.

[33] Nielsen MB. Model systems for understanding absorption tuning by opsin proteins. Chemical Society reviews. 2009;38:913-24.

[34] Nakanishi K, Crouch R. Application of artificial pigments to structure determination and study of photoinduced transformations of retinal proteins. Israel J Chem. 1995;35:253-72.

[35] Houjou H, Inoue Y, Sakurai M. Study of the opsin shift of bacteriorhodopsin: Insight from QM/MM calculations with electronic polarization effects of the protein environment. J Phys Chem B. 2001;105:867-79.

[36] Melaccio F, Ferre N, Olivucci M. Quantum chemical modeling of rhodopsin mutants displaying switchable colors. Physical Chemistry Chemical Physics : PCCP. 2012;14:12485-95.

[37] Coto PB, Strambi A, Ferre N, Olivucci M. The color of rhodopsins at the ab initio multiconfigurational perturbation theory resolution. Proceedings of the National Academy of Sciences of the United States of America. 2006;103:17154-9.

[38] Wang WJ, Nossoni Z, Berbasova T, Watson CT, Yapici I, Lee KSS, et al. Tuning the Electronic Absorption of Protein-Embedded All-trans-Retinal. Science. 2012;338:1340-3.

[39] Butt HJ, Fendler K, Bamberg E, Tittor J, Oesterhelt D. Aspartic Acid-96 and Aspartic Acid-85 Play a Central Role in the Function of Bacteriorhodopsin as a Proton Pump. EMBO J. 1989;8:1657-63.

[40] Hirai T, Subramaniam S. Structural insights into the mechanism of proton pumping by bacteriorhodopsin. FEBS Lett. 2003;545:2-8.

[41] White SN, Chave AD, Reynolds GT. Investigations of ambient light emission at deep-sea hydrothermal vents. J Geophys Res-Sol Ea. 2002;107.

[42] Ludlam CF, Sonar S, Lee CP, Coleman M, Herzfeld J, RajBhandary UL, et al. Site-directed isotope labeling and ATR-FTIR difference spectroscopy of bacteriorhodopsin: the peptide carbonyl group of Tyr 185 is structurally active during the bR-->N transition. Biochemistry-US. 1995;34:2-6.

[43] Luecke H, Schobert B, Richter HT, Cartailler JP, Lanyi JK. Structural changes in bacteriorhodopsin during ion transport at 2 Angstrom resolution. Science. 1999;286:255-60.

[44] Calimet N, Ullmann GM. The influence of a transmembrane pH gradient on protonation

probabilities of bacteriorhodopsin: the structural basis of the back-pressure effect. Journal of Molecular Biology. 2004;339:571-89.

[45] Imasheva ES, Balashov SP, Choi AR, Jung KH, Lanyi JK. Reconstitution of Gloeobacter violaceus rhodopsin with a light-harvesting carotenoid antenna. Biochemistry-US. 2009;48:10948-55.

[46] Balashov SP, Imasheva ES, Boichenko VA, Anton J, Wang JM, Lanyi JK. Xanthorhodopsin: a proton pump with a light-harvesting carotenoid antenna. Science. 2005;309:2061-4.

[47] Luecke H, Schobert B, Stagno J, Imasheva ES, Wang JM, Balashov SP, et al. Crystallographic structure of xanthorhodopsin, the light-driven proton pump with a dual chromophore. Proceedings of the National Academy of Sciences of the United States of America. 2008;105:16561-5.

[48] Kralj JM, Douglass AD, Hochbaum DR, Maclaurin D, Cohen AE. Optical recording of action potentials in mammalian neurons using a microbial rhodopsin. Nature Methods. 2012;9:90-U130.

[49] Peteanu LA, Schoenlein RW, Wang Q, Mathies RA, Shank CV. The first step in vision occurs in femtoseconds: complete blue and red spectral studies. Proceedings of the National Academy of Sciences of the United States of America. 1993;90:11762-6.

[50] Polli D, Altoe P, Weingart O, Spillane KM, Manzoni C, Brida D, et al. Conical intersection dynamics of the primary photoisomerization event in vision. Nature. 2010;467:440-3.

[51] Laricheva EN, Gozem S, Rinaldi S, Melaccio F, Valentini A, Olivucci M. Origin of Fluorescence in 11-cis Locked Bovine Rhodopsin. J Chem Theory Comput. 2012;8:2559-63.

[52] Otto H, Marti T, Holz M, Mogi T, Stern LJ, Engel F, et al. Substitution of amino acids Asp-85, Asp-212, and Arg-82 in bacteriorhodopsin affects the proton release phase of the pump and the pK of the Schiff base. Proceedings of the National Academy of Sciences of the United States of America. 1990;87:1018-22.

[53] Saeedi P, Moosaabadi JM, Sebtahmadi SS, Behmanesh M, Mehrabadi JF. Site-directed mutagenesis in bacteriorhodopsin mutants and their characterization for bioelectrical and biotechnological equipment. Biotechnology Letters. 2012;34:455-62.

[54] Saeedi P, Moosaabadi JM, Sebtahmadi SS, Mehrabadi JF, Behmanesh M, Nejad HR, et al. Generation and analysis of bacteriorhodopsin mutants with the potential for biotechnological applications. Bioengineered. 2012;3:275-9.

[55] Tittor J, Haupts U, Haupts C, Oesterhelt D, Becker A, Bamberg E. Chloride and proton transport in bacteriorhodopsin mutant D85T: different modes of ion translocation in a retinal protein. Journal of Molecular Biology. 1997;271:405-16.

[56] Briand J, Leonard J, Haacke S. Ultrafast photo-induced reaction dynamics in bacteriorhodopsin and its Trp mutants. J Optics-UK. 2010;12.

[57] Marti T, Otto H, Mogi T, Rosselet SJ, Heyn MP, Khorana HG. Bacteriorhodopsin Mutants Containing Single Substitutions of Serine or Threonine Residues Are All Active in Proton Translocation. Journal of Biological Chemistry. 1991;266:6919-27.

[58] Rothschild KJ, He YW, Sonar S, Marti T, Khorana HG. Vibrational Spectroscopy of Bacteriorhodopsin Mutants - Evidence That Thr-46 and Thr-89 Form Part of a Transient Network of Hydrogen-Bonds. Journal of Biological Chemistry. 1992;267:1615-22.

[59] Russell TS, Coleman M, Rath P, Nilsson A, Rothschild KJ. Threonine-89 participates in the active site of bacteriorhodopsin: Evidence for a role in color regulation and Schiff base proton transfer. Biochemistry-US. 1997;36:7490-7.

[60] Mogi T, Marti T, Khorana HG. Structure-function studies on bacteriorhodopsin. IX. Substitutions of tryptophan residues affect protein-retinal interactions in bacteriorhodopsin. The Journal of Biological Chemistry. 1989;264:14197-201.

[61] Peralvarez-Marin A, Marquez M, Bourdelande JL, Querol E, Padros E. Thr-90 plays a vital role in the structure and function of bacteriorhodopsin. The Journal of Biological Chemistry. 2004;279:16403-9.

[62] Subramaniam S, Greenhalgh DA, Rath P, Rothschild KJ, Khorana HG. Replacement of leucine-93 by alanine or threonine slows down the decay of the N and O intermediates in the photocycle of bacteriorhodopsin: implications for proton uptake and 13-cis-retinal----all-trans-retinal reisomerization. Proceedings of the National Academy of Sciences of the United States of America. 1991;88:6873-7.

[63] Greenhalgh DA, Farrens DL, Subramaniam S, Khorana HG. Hydrophobic amino acids in the retinal-binding pocket of bacteriorhodopsin. The Journal of Biological Chemistry. 1993;268:20305-11.

[64] Jang DJ, el-Sayed MA, Stern LJ, Mogi T, Khorana HG. Effect of genetic modification of tyrosine-185 on the proton pump and the blue-to-purple transition in bacteriorhodopsin. Proceedings of the National Academy of Sciences of the United States of America. 1990;87:4103-7.

[65] Peralvarez-Marin A, Bourdelande JL, Querol E, Padros E. The role of proline residues in the dynamics of transmembrane helices: the case of bacteriorhodopsin. Molecular Membrane Biology. 2006;23:127-35.

[66] Ahl PL, Stern LJ, During D, Mogi T, Khorana HG, Rothschild KJ. Effects of amino acid substitutions in the F helix of bacteriorhodopsin. Low temperature ultraviolet/visible difference spectroscopy. The Journal of Biological Chemistry. 1988;263:13594-601.

[67] Hashimoto S, Obata K, Takeuchi H, Needleman R, Lanyi JK. Ultraviolet resonance Raman spectra of Trp-182 and Trp-189 in bacteriorhodopsin: novel information on the structure of Trp-182 and its steric interaction with retinal. Biochemistry-US. 1997;36:11583-90.

[68] Subramaniam S, Marti T, Khorana HG. Protonation state of Asp (Glu)-85 regulates the purple-to-blue transition in bacteriorhodopsin mutants Arg-82----Ala and Asp-85----Glu: the blue form is inactive in proton translocation. Proceedings of the National Academy of Sciences of the United States of America. 1990;87:1013-7.

[69] Wang WW, Sineshchekov OA, Spudich EN, Spudich JL. Spectroscopic and photochemical characterization of a deep ocean proteorhodopsin. The Journal of Biological Chemistry. 2003;278:33985-91.

[70] Maiti TK, Yamada K, Inoue K, Kandori H. L105K mutant of proteorhodopsin. Biochemistry-US. 2012;51:3198-204.

[71] Tamogami J, Kikukawa T, Nara T, Shimono K, Demura M, Kamo N. Photoinduced proton release in proteorhodopsin at low pH: the possibility of a decrease in the pK(a) of Asp227. Biochemistry-US. 2012;51:9290-301.

[72] Black SD, Mould DR. Development of Hydrophobicity Parameters to Analyze Proteins Which Bear Posttranslational or Cotranslational Modifications. Anal Biochem. 1991;193:72-82.

[73] Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods. 2009;6:343-5.

[74] Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. Nature Protocols. 2009;4:1-13.

[75] Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 2006;22:195-201.

[76] Bosley AD, Ostermeier M. Mathematical expressions useful in the construction, description and evaluation of protein libraries. Biomolecular Engineering. 2005;22:57-61.

Figure Legends

Figure 1: Schematic of screen for GR spectral tuning. Briefly, mutant GR libraries were transformed into *E. coli* followed by growth on selective solid media. Single *E. coli* colonies were picked with toothpicks, and grown in 96-well plates. Following induction, His-tagged mutant GRs were purified using Ni-NTA resin. Absorbance spectra of mutant GRs were measured with a standard plate reader, from which the absorption maxima were determined.

Figure 2: Retinal binding pocket in GR homology model. Side chains of amino acids within 5 Å of retinal are shown. Glycines are displayed with hydrogen atoms visible. Residues for which only blue shifts were detected are colored blue (W122, V126, M158, G162, E166, G178, S181), residues with only red shifts are colored red (T125, L129, W222, Y225, P226, D253), and residues that yield both blue and red shifts are colored yellow (D121, A256). Black indicates positions for which no shifts were identified (I159, T182, F185, Y229) and the conserved Schiff base residue (K257), which was not targeted for mutagenesis. The retinal molecule is shown in orange. Helix 6 is omitted for clarity.

Figure 3: Spectrally-shifted variants identified by site-saturation mutagenesis and recombination. The absorption maxima (λ_{max}) represent averages, with standard error, of three biological replicas that were grown and extracted separately. Shifts are reported with respect to λ_{max} of WT GR (538 nm). H^+-pumping activities were determined using a dye-efflux assay (see Methods): (-) no detectable activity, (+) activity that is less than 50% of WT, (++) activity that is more than 50% of WT. Dashed lines separate mutants resulting from different rounds of directed evolution. From the top, these are: 1) site-saturation mutagenesis of WT GR, 2) recombination libraries, and 3) second round of site-saturation mutagenesis, performed on G162L/E166W/A256S and T125C/A256M. New mutations are shown in bold and parent mutations in plain text. (insert) Purified variants of GR, including the wild type protein (GR WT), 3 GR variants with red-shifted λ_{max} (GRr1, GRr2, and GRr3), and 3 GR variants with red-shifted λ_{max} (GRb1, GRb2, and GRb3). GRr1 = T125N, GRr2 = T125C/A256M, GRr3 = D121E/T125C/A225M, GRb1 = G162L, GRb2 = G162L/E166W/A256S, and GRb3 = W122L/G162L/E166W/A256S.

Figure 4: Relationship between hydrophobicity and residue distance from the Schiff base (K257) for blue- and red-shifting mutations. Amino acid substitutions that are more hydrophobic than the parental residues result in ΔHydrophobicity > 0, while those that are less hydrophobic than the parental residues have ΔHydrophobicity < 0. For blue tuning mutations, the pearson correlation between ΔHydrophobicity and distance from K257 is 0.55 (p-value = 0.0045), while for red tuning mutations, there is no correlation. Amino acid hydrophobicity scores were obtained from Black and Mould [72].

Figure 5: Normalized absorption spectra of WT GR and key variants from each round of directed evolution. Absorption spectra are colored to match the colors of the rhodopsin pigments.

Figure 6: Pairwise comparison of proton pumping and distance of mutation away from K257. GR single mutants are binned based on their proton-pumping activity. Box-and-whisker plot shows the relationship between activity and the distance of mutated residue from K257. Black lines inside the boxes indicate the second quartile (the median value), box left and right bound indicate the first and third quartile, and the whiskers indicate the lowest and highest values. The raw data are overlaid as individual points and colored according to their spectral shifts. The total number of data points per category is 15 (++), 17 (+), and 20 (-).

Figure 7: Box-and-whisker plot of spectral shifts of variants indicating their proton-pumping activities according to: (-) no detectable activity, (+) activity less than 50% of WT, and (++) activity that is more than 50% of wild-type. For blue shifts there are 10 (++), 11 (+), and 4 (-) data points, while for red shifts there are 5 (++), 6 (+), and 16 (-) data points. Black lines inside the boxes indicate the median value, box left and right bound indicate the first and third quartile, and the whiskers indicate the lowest and highest values (except outliers, which are shown individually as circles).

Figure 8: Fluorescent properties of GR variants measured in live *E. coli*. (A) The area under curve (AUC) was computed from the excitation spectra of 70 unique GR, ranked-ordered, and color-coded according to spectral shift (technical replicates of GR(D121N), a red-shifted variant, are shown in green). (B) Spline-fit excitation/emission spectra (performed in biological quadruplicate) of 6 outliers from (A) and WT GR. The mutations are written in shorthand without the residue numbers: GR(DETCAM) = GR(D121E/T125C/A256M), GR(TCFTAM) = GR(T125C/F185T/A256M), GR(TCAM) = GR(T125C/A256M), GR(TVAM) = GR(T125V/A256M), GR(TCWMAM) = GR(T125C/W222M/A256M), GR(DA) = GR(D121A). (C) Fluorescence of GR(D121E/T125C/A256M) normalized for total expression measured from pH 6-11.

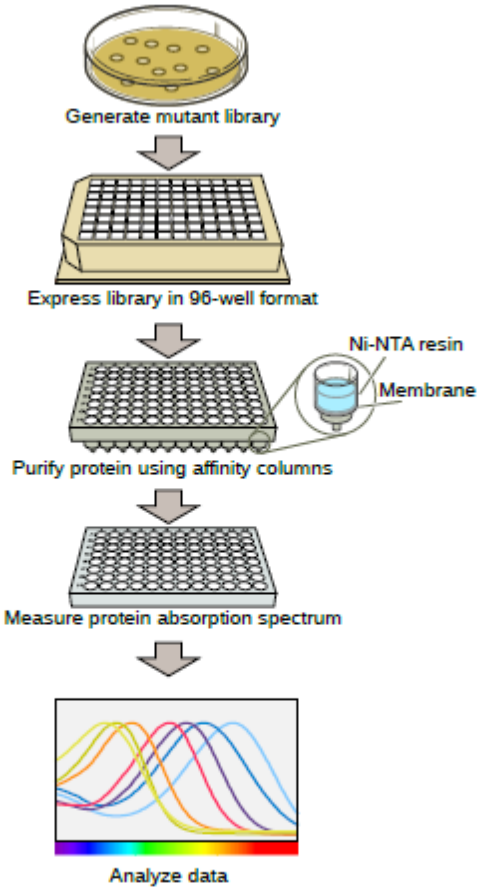


Figure 1

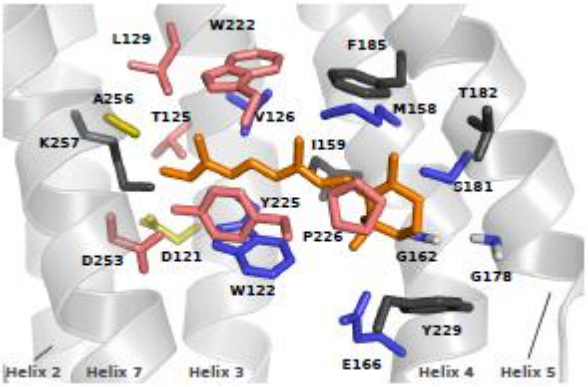


Figure 2

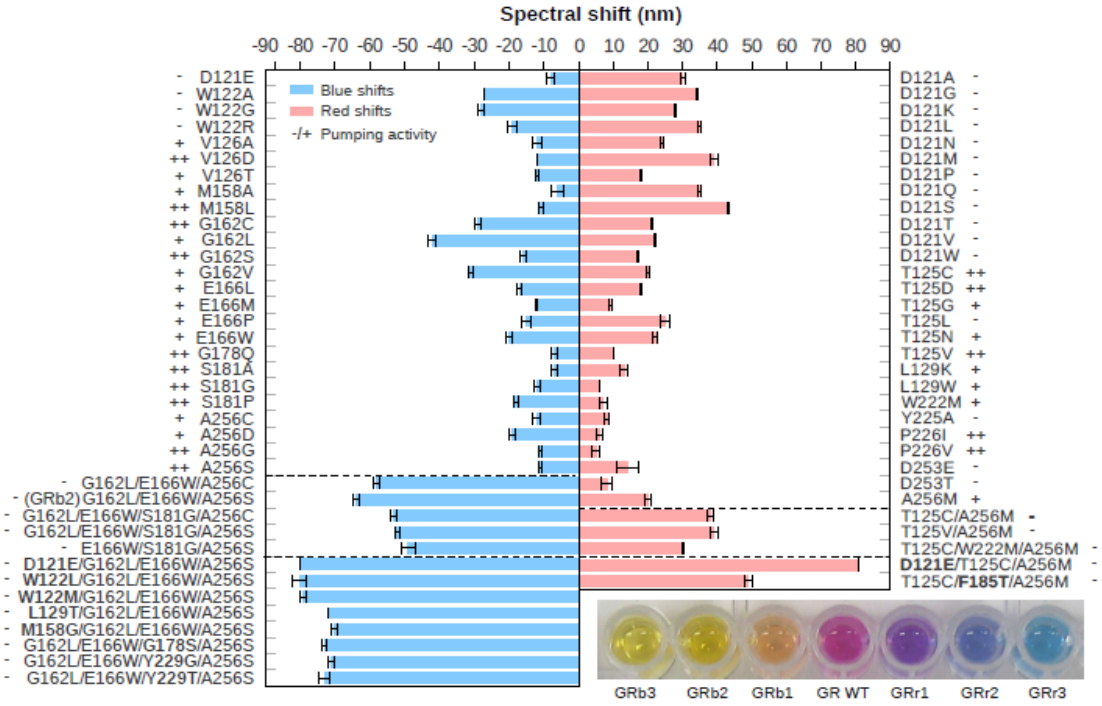


Figure 3

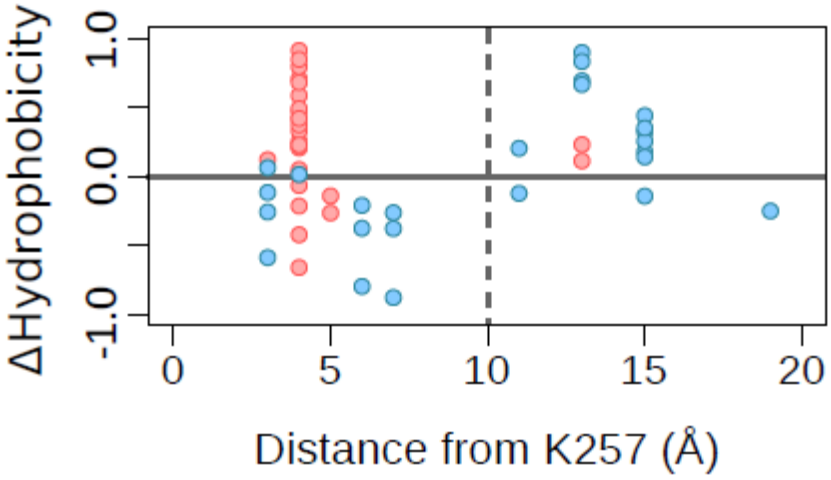


Figure 4

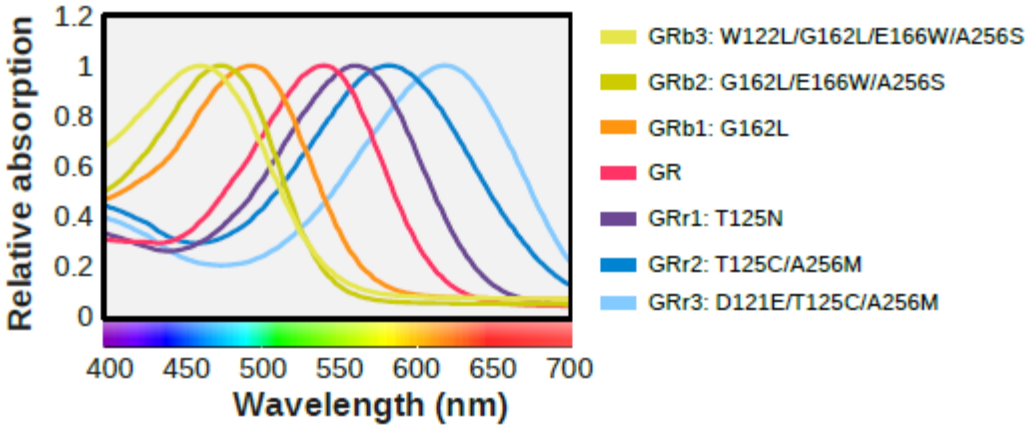


Figure 5

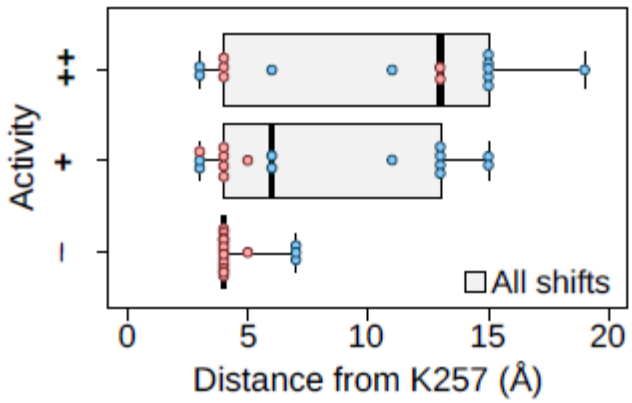


Figure 6

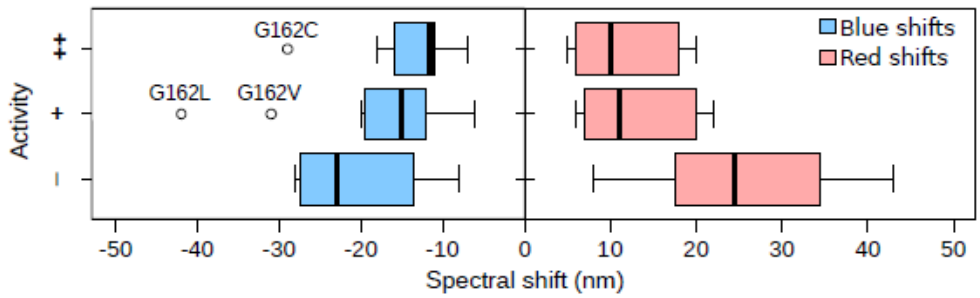


Figure 7

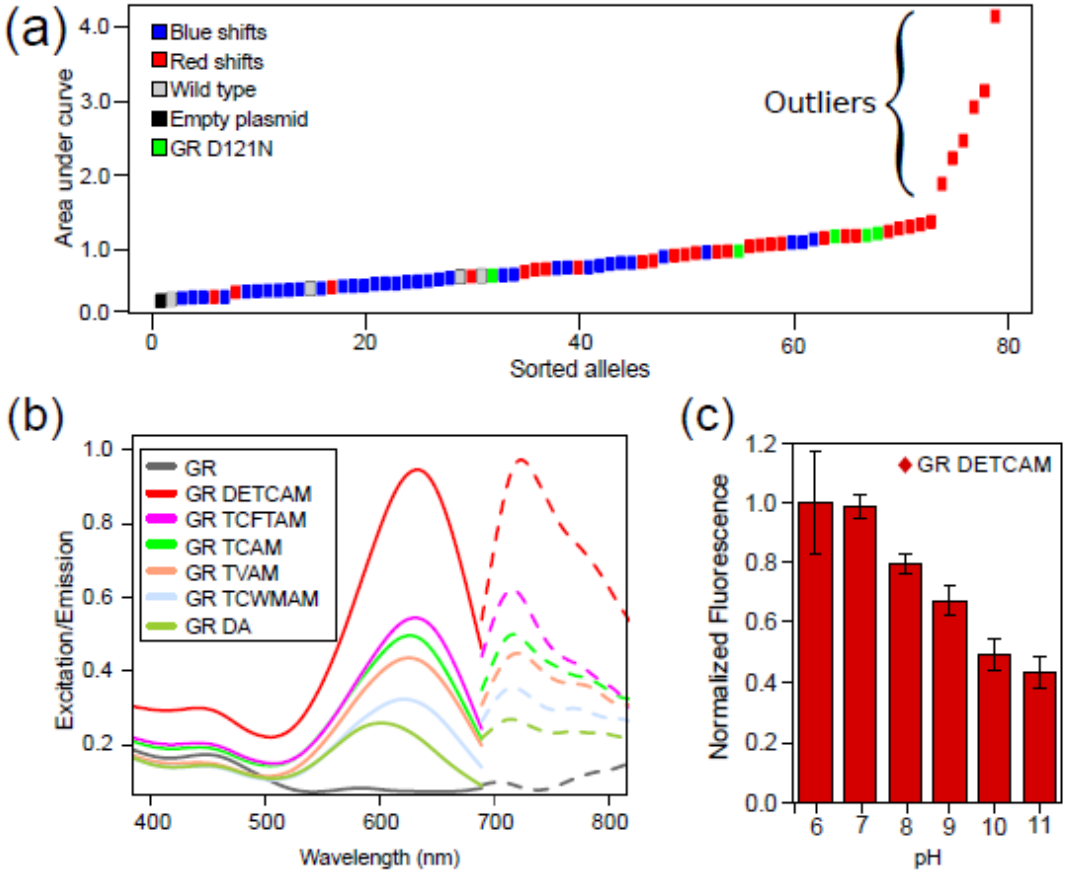
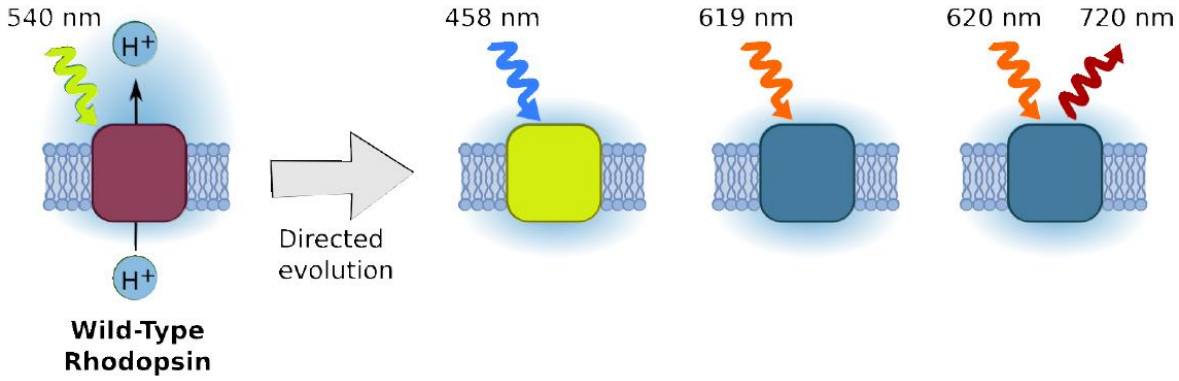


Figure 8



Graphical abstract

Tables

Table 1. Comparison of natural variation and introduced mutations in retinal-binding pocket of GR and homologs.

GR Position[a]	121	122	125	126	129	158	159	162	166	178	181	182	185	222	225	226	229	253	256	257
Natural variation[b]	D	W	T	V	L	M	I L V	G	E D	G F W	S A	T M S	F	W	Y	P	Y F	D	A	K
Mutations[c]	A, P E, Q G, S K, T L, V N, W M	A G R	C D G L N V	A D T	K W	A L	-	C L S V	L M P W	Q	A G P	-	-	M	A	I V	-	E T	C D G M S	-

BR Position[a]	85	86	89	90	93	118	119	122	126	138	141	142	145	182	185	186	189	212	215	216
Natural variation[b]	D	W	T	T	L	M	I V	G	A T	W	S	D T	F L M	W	Y	P	W	D	A S	K
Mutations[c]	A[52] E[53] N[53] V[54] S[54] T[55]	F[56]	A[57] D[58] N[59] V[60]	A[61] V[61]	A[62] T[62] V[62]	A[26]	-	C[63]	-	C[60]	A[26] C[57]	V[57]	A[63]	F[56]	F[64]	A[65] L[66] V[60]	F[67]	E[68]	N[26] T[26]	-

PR Position[a]	97	98	101	102	105	133	134	137	141	151	154	155	158	197	200	201	204	227	230	231
Natural variation[b]	D	W I	T P Q	V C	L Q A V Y E	V A I L	M	A F G V	G	A C G	I L V	A G	A G	W	Y	P	Y	D	N	K
Mutations[c]	N[69]	G[25]	A[25]	A[25]	K[70]	-	-	-	-	-	-	-	-	-	H[25] N[25]	-	-	E[71]	-	-

[a] Homologous residues in the retinal-binding pocket were identified through structure-guided alignment of protein sequences. [b] Natural variation represents variants retrieved from BLAST searches of the NCBI database with GR, BR and PR. [c] Mutations conferring spectral tuning in each of the three proteins, and close homologs. The coloring of the mutations indicates whether they are blue- or red-shifted. For GR the 52 single mutations were discovered in the first round of site-saturation mutagenesis. For BR and PR mutations were identified through a literature search. GR mutations that correspond to mutations previously described in BR or PR are underlined.

Table 2. Quantum yields of bright GR variants identified in this study.

Mutant	Quantum Yield
D121A	5×10^{-3}
T125C/W222M/A256M	8×10^{-3}
T125C/A256M	9×10^{-3}
T125V/A256M	9×10^{-3}
T125C/F185T/A256M	9×10^{-3}
D121E/T125C/A256M	1.2×10^{-2}

Highlights

- Performed mutagenesis of the retinal-binding pocket of a proton-pumping rhodopsin.
- Shifts in absorption maxima of up to +/- 80 nm were achieved.
- A subset of mutants exhibited strong fluorescence in the far-red/infrared.
- A first step towards brighter opsin-based biological sensors.